DOCUMENT RESUME

ED 411 277 TM 027 260

AUTHOR Farr, Beverly P.; Garcia-Sims, Angela; Garcia Fontana, Rose Marie

TITLE The Complexities of Evaluating Bilingual Education Programs.

PUB DATE 1997-03-00

NOTE 14p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).

PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Access to Education; *Bilingual Education; Cultural Differences; Educational Assessment; Elementary Secondary Education; *Equal Education; Evaluation Methods; Language Minorities; Limited English Speaking; Minority Groups; *Performance Based Assessment; Spanish Speaking; *Special Needs Students; *Standards; Test Bias; Testing Problems

IDENTIFIERS Arizona Student Assessment Program

ABSTRACT

As part of a discussion of the complexities of evaluating bilingual education, this paper discusses critical questions related to equitable education. A first concern is the development of standards. Children must be given equal opportunities to learn, but the development of standards and new assessments alone will not ensure that inequities will disappear. Performance assessments are not an answer either, since the implementation of performance assessments will show that test bias does not simply go away. Potential causes of bias in performance assessments will be language-dependent skills and the problem of situational or cultural contexts in the assessments that are unfamiliar to students from certain cultures. Practical approaches to ensuring accuracy in assessments for students with diverse learning needs begin with providing reasonable access to manipulatives and resources when possible and making sure the language of the directions is clear. Educators should request special administration conditions for students when needed. An example of mediated administration that is more equitable for students is presented in the Arizona Student Assessment Program (ASAP) (no longer the state assessment) and its accommodations for Spanish speaking students. Strategies are included for administering the ASAP in English to limited English proficient students. (Contains two tables.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

The Complexities of Evaluating Bilingual Education Programs

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Beverly Farr
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

AERA Roundtable Session 29.04, 3:05-3:45 p.m.
Wednesday, March 26, 1997
Sheraton Ballroom II, Level 4

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Beverly P. Farr, Angela García-Sims, Rose Marie García Fontana
WestEd, Northern California Comprehensive Assistance Center

ED 411 277

AGENDA

I. **Introductions/descriptions of our individual experiences with Bilingual Education Programs**

II. **Ideal Program Evaluation Process and Components**

III. **Phases of an Evaluation/Complexities in Evaluating Bilingual Education Programs**

IV. **Lessons learned/Recommendations**

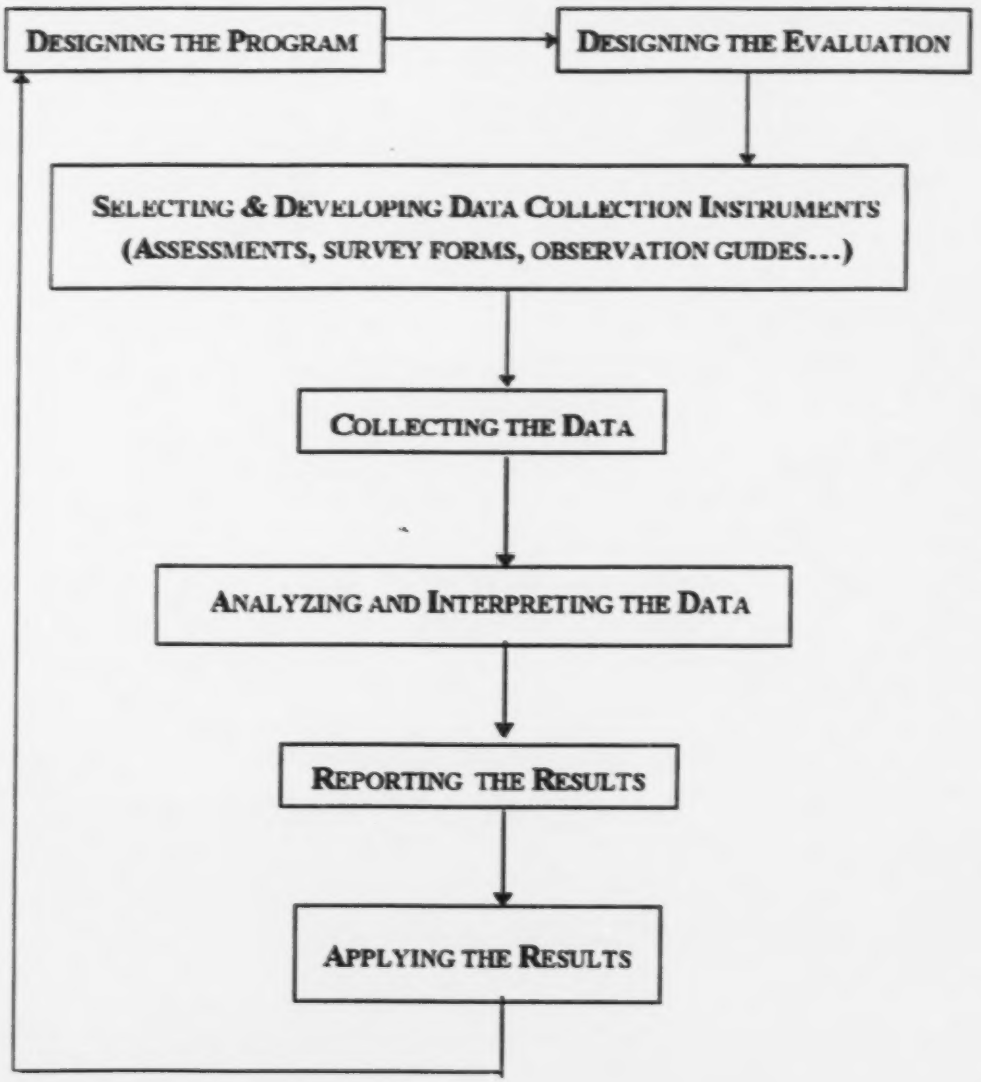
V. **Participants share experiences with Bilingual Education Evaluation: challenges and solutions**

VI. **Closing**

TM027260

BEST COPY AVAILABLE

EVALUATION PHASES



ANGELA GARCIA-SIMS, WESTED CAC

Factors Complicating Bilingual Education Program Evaluations

Angela Garcia-Sims, Ph.D.

Complicating Factor	Possible Resolutions
1. **Language and languages in contact:** Some students have little proficiency in any language. They present the challenges of any student with low language ability complicated by the sociolinguistics of language contact	1. Assessments in the students' primary and second languages. (See companion paper by Beverly Farr). In addition to student assessment instruments, bilingual evaluations must ensure language access in any parent or community surveys and other data collection from potential non-English speakers
2. **Culture** – the sociocultural context of the program, e.g., Parent attitudes toward second language learning and retention of the first language influence their children's performance on assessments as well as their response to instruction any specific language.	2. Awareness of the sociolinguistic setting; survey parents and students on attitudes toward languages and schooling, goals for students. Include interpretation of evaluation data.
3. **Politics** Emphatic factions continue to attack bilingual education and any governmental support of instruction in languages other than English. These continue to lobby both in and out of the media for elimination of any bilingual program support.	3. Include as all major stakeholders in the development and conduct of the evaluation; acknowledge and include political awareness in evaluation design; provide interim reports to stakeholders and local media; prepare clients for potential controversies. Include external, detached observers-evaluators
4. **Pedagogical Theory**—specifically second language learning and instruction. Strong voices both in and outside the education community object to instructional programs that include primary language instruction asserting that such instruction delays acquisition of English and promotes disunity among different nationalities.	4. Include evaluation questions and collect data to examine alternative explanations for program outcomes; provide information on research comparing different models (e.g., Collier & Johnson, Baker et. al). Examine long-term effects of high academic achievement and school success. (See companion paper by R.M. Fontana)

BEST COPY AVAILABLE

Critical Questions related to Equitable Assessment

As stated by Shirley Malcom, the goal must be to "reinvent assessment to support equity;" this will require us to answer some critical questions.

Is the development of standards the place to start?

Educators who have disparaged standards as the appropriate starting place for educational reform seem generally committed to correcting basic inequalities in the schools as the first priority. Some of these critics are among those who have written forcefully about the need for educational reform or have themselves been architects of school reform models that have been implemented with a considerable degree of success. James Comer, for example, believes that "what poor children need first are the social skills and community bonds that children of privilege take for granted." Similarly, the Coalition of Essential Schools' founder, Ted Sizer, warns that racing ahead with a plan to establish standards and tests before basic inequalities among schools–class size, outmoded facilities, inequitable school financing–are addressed is "the wrong sequence. We will end up proving once again that poor kids don't score as well as rich kids" (O'Neil, 1993). Jonathan Kozol likewise argues that the first priority must be to eliminate separate and unequal school systems. He believes this goal is attainable by paying the highest salaries to the best teachers who serve the poorest children in classrooms with a class size of no more than 20 in urban schools.

Will all children be given equal opportunities to learn?

This concern is certainly related to the preceding one, but it is worth mentioning, because those who are developing standards must remember that if they are using that process as the starting place for reform, they must maintain a parallel focus on standards that ensure access to learning. What were initially referred to as "school delivery standards" are now more commonly referred to as "opportunity to learn" standards and are intended to provide a basis for judging whether schools are providing students with sufficient opportunity to attain high standards. Absent such assurance, it is easy to understand how unfair it would be to impose standards on students who do not have access to the teaching and resources necessary to ensure the highest level of learning.

Will the development of standards and new assessments, alone, ensure that these inequities will disappear?

Of course not. Concerted attention to the development of opportunity to learn standards may provide needed protection, but educators who have begun to grapple with the difficult problem of equity are finding it tough going. As the number of schools across the nation implementing powerful instructional programs and innovative assessment strategies continues to increase, it will become easier to define the opportunities to learn that must be available to all students. This could be a long journey, however, and it would seem best to embark on the path, outfitted with the wealth of information now available about how children learn–including children who may be learning English or who may be from a cultural group different from the "mainstream".

Will performance assessments make a difference for children who are traditionally low-performing?

This is a crucial point, because many critics of standardized, norm-referenced tests who are now advocating for performance-based assessments believe that these new assessments will provide a richer, more complete picture of what children have learned. But, as noted earlier in a comment by Sizer, we may, once again, merely emphasize the performance gap between poor kids and rich kids. While many believe that the new assessments will eventually present the greatest potential for learning about the learning of our diverse student population, it is folly to presume that low-performing students–often those from linguistic and cultural minorities–will suddenly do better on such assessments, or, at least, that we will (potentially) get a truer picture of their learning. The latter is likely, but, in fact, initial studies seem to indicate that the use of such assessments may widen the gap between disadvantaged students and their more advantaged counterparts. While these studies are clearly preliminary, it is not difficult to see why this might be the case.

Lorrie Shepard (1993) has conducted several studies to evaluate differences in performance among student groups. Based on a review of her own studies and others conducted in Great Britain, she concludes that the results refute the assumption that because performance assessments are more valid, there will be fewer differences in performance. A widened gap seems, in fact, to have materialized for most minorities, with the exception of females, who appear to do better with open-ended questions. Shepard offers two possible explanations for this outcome: biased measurement and lack of opportunity to learn.

Test bias does not simply go away with the design and implementation of performance-based assessments. Potential causes of test bias include a greater reliance on language-dependent skills: more reading and writing is required for extended problems and such activities as writing explanations for mathematical solutions. In addition, performance assessments are often based on situational contexts or cultural items that may be unfamiliar to students from certain cultures. The critical thinking demands typical of such assessments, such as making judgments or expressing values, may not be compatible with some cultures' norms.

General Guidelines for Developing Equitable Assessments

As noted by Malcom (1991), "Equity concerns must be at the heart of assessment design." The conditions that Malcom offers are fairly simple and straightforward. They will serve as a framework for our discussion of issues related to the design and use of equitable assessments.

1. Rules about what is to be known must be clear to all.
2. Resources needed to achieve must be available to all.
3. Ways of demonstrating knowledge must be many and varied.
4. Things valued by different groups must be reflected in the original statement of what is to be known.

(Excerpted from Farr & Trumbull, *Assessment Alternatives for Diverse Classrooms*. Norwood, MA: Christopher-Gordon, 1997.)

BEST COPY AVAILABLE

Ways to Reduce Bias in Assessment of Cultural Minority/ELL Students

- Use multiple measures at multiple points in time. Never use a single test to make a major decision about a student.
- Allow flexibility in administration and in types of tasks
- Consider background information (context) in interpreting student performance. Low scores cannot be considered valid for ELL students.
- Draw on knowledge of community to understand learning styles, problem-solving approaches, ways of using language. (adds to context knowledge)
- Consider social factors that may inhibit best possible performances.
- Moderate student work. That is, have teachers get together regularly to review student work so as to "stay on same wavelength" about what is valued and what counts as evidence for student success.

Benefits of Portfolios for Instruction and Assessment

- Are well matched to a "thinking curriculum" (active, collaborative)
- Involve students in process of self-evaluation and reflection
- Motivate student ownership, sense of accomplishment
- Convey importance of student work
- Help teachers see how students think
- Help students and teachers set goals
- Provide developmental record of student achievement
- Have no "floor" or "ceiling" i.e., start and finish where the student is
- (potentially) Show relationships among various skills, subject areas
- Directly inform instruction
- Serve as a good communication tool with parents
- Provide rich context in which to understand other kinds of (test) performances

A Few Suggestions...

- Consider *up front* how any change in assessment policy will affect LEP, bilingual, bicultural, migrant, special education/learning handicapped, gifted students.
- Be particularly cautious about making inferences from assessments administered in a student's second language.
- Never rely on a single test/assessment score or outcome to determine achievement level or proficiency of "special population" students.
- Support instruction and assessment that draw on students' personal experience, demand use of multiple intelligences, are meaningful to students.
- Accommodate (when possible) student differences in rate of processing, need for additional time to complete tasks.
- Re-examine your own beliefs about *who* can learn. Expectations are powerful.
- Remember that giving/taking a standardized test is itself a communication event that calls for use of particular language skills. Middle class children's home uses of [decontextualized] language prepare them better for test-taking.

Cautions...

- Performance tasks -- the kind that tend to go into portfolios--may be more difficult for some students than traditional tests.
- Portfolios demand a high level of student participation, something that will take time and teacher support for many students to demonstrate.
- Portfolio entries typically make a high demand on language skill. Considerations need to be made for bilingual and learning handicapped students.
- Interpretations of student performances can be tricky, especially if mediated or modified by teacher. Criteria must be clearly established.
- Teachers need to coordinate with each other; it is not good to have multiple portfolio systems operating within one district or collaborative.
- Student or teacher mobility can introduce discontinuities in a portfolio system.

BEST COPY AVAILABLE

Some practical matters...
Ensuring accuracy in assessments for students with diverse learning needs

1. Provide reasonable access to manipulatives and resources throughout the test, except when items are explicitly measuring that specific skill (such as the use of a calculator or dictionary).

2. Make sure the language of the items and directions is clear.
   - Give the test to a sample of students and their teachers and have them circle confusing words and phrases and suggest solutions. Suggest changes as necessary.
   - If possible, apply plain language principles to the items
   - Convene a small group of content experts and the test contractor/vendor, ask if the changes you have suggested change the nature of the content which is to be tested. If the integrity of the content remains unchanged, make the changes.

3. Make sure the rubrics from open-ended questions are accurately measuring a full range of responses.
   - Review student work and the training packets, and ask the following questions per item:
     ⇒ Does the rubric allow for diverse types of responses, including (1) non-text format responses, (pictures, diagrams, charts, algorithms, etc.), and (2) reasonable responses of students from "non-standard" backgrounds?
     ⇒ What is the % of items which allow for these diverse responses? -- How many?
     ⇒ As needed, supplement the training packets with diverse responses and request forms with a reasonable percentage of satisfactory rubrics.

4. Request oral administration and/or response options for students with diverse needs, including but not limited to, special ed and LEP students (L1 as well as L2). These are to be aggregated with text administered/scored responses.
   - Why not? Think standards based -- how accurate are the scores if the student can't read the test (even if they draw pictures with the bubbles)?
   - For test administration, read aloud (although cueing is a problem), use cassette tapes, video, computer with voice.
   - Oral response includes use of a scribe (although accurate recording is sometimes a problem and it is expensive), cassette, or voice-receiving computer.
   - Issue with cassette recorder and oral output from computer is cost of scoring in real time.
   - Look for trained volunteers to score under same controlled conditions (who have same qualifications as those hired by test publishers), $ from other programs to support additional contractor scoring.
   - Technology is coming on line to permit printed output from voice recorded computer input. This can be scored at the same cost as text responses.

5. Request other administration options for students with diverse needs (including but not limited to....).
   - Options include setting, time, less items per page and larger print, breaks, templates.
   - Why not? This should be ok if the option is not related to what is being measured.

BEST COPY AVAILABLE

- Use parent volunteers, other community members to help with additional staffing.
- Staff can prepare forms which are large print and less items per page. These can be locally copied under direction from the contractor.

6. Push for tests which include performance components-- some kind of activity with associated items, in addition to paper and pencil sections. These don't have to be lengthy.
   - Broadens points of access
   - We know we measure writing abilities better by having students actually write. Don't downplay this in other disciplines -- e.g. students should DO science, not just respond to science someone else does.
   - False that only "lower" students benefit from hands-on or discussion experiences. Get evidence from the local gifted and talented expert.

(Adapted from Rebecca Kopriva, Ph.D., CCSSO presentation)

An Example of Mediated Administration

Arizona Student Assessment Program

(Please note that the ASAP is no longer the state assessment in Arizona. The guidelines presented are merely offered as an example.)

Guidelines

Eligibility for Spanish and Mediated English

There are no exemptions from the ASAP for limited English proficient students. All students will take the ASAP in English without mediation except for students who meet the following criteria:

ASAP Spanish Assessment:

- Students who are enrolled in a Spanish bilingual program and are receiving their primary instruction in Spanish

OR

- Students who are enrolled in the first year of an English as a Second Language (ESL) program in Arizona and their recent, previous schooling was in a setting where Spanish was the primary language of instruction.

Mediated English ASAP:

- Students who are enrolled in the first year of an English as a Second Language program and who would experience difficulty taking the English version of the ASAP due to limited proficiency in English.

OR

- Students who are enrolled in a bilingual program where the primary language is other than Spanish and the students are receiving reading and writing instruction in their native language.

Mediated Spanish ASAP:
Students who are enrolled in the first year of a bilingual or ESL program may take a mediated Spanish ASAP if the student is unable to take the assessment without this assistance.

- Students who meet the criteria for the Spanish ASAP or the mediated English are eligible to take those forms of the ASAP. However, they are not required to take them. They may take the unmediated English version if the teacher decides this is best for the student.

Guidelines for Mediated ASAP

Teachers may assist students who are eligible to take the mediated English ASAP in any way that is necessary provided they do not give the answers or lead students to answers.

Strategies include but are not limited to:

- flexible scheduling
- flexible setting
- translation
- simplification of language
- reading to student
- taking dictation
- use of visuals

Strategies for Administering the ASAP in English to *LEP Students

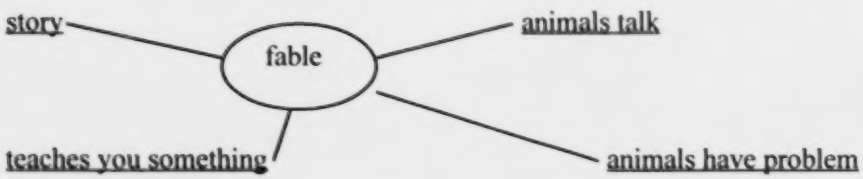
Directions only:

You do not have to read the directions word for word. Simplify! You may:
—translate any portion or all of the directions
—translate the title of the story and key words that appear in the directions (if you are not a speaker of the language, ask someone ahead of time & write on board)
—clarify and explain difficult terms used in the directions and put simple definitions on the board or chart paper

e.g. character= person or animal in a story
moral = teaches a lesson
agree = that's what I think, too

—simplify/rephrase difficult sentences or phrases (abstract or sophisticated language)
e.g. My response states the moral of the story. = My answer tells what the moral is.

—use visuals:
real objects
illustrations/charts
pictures (ant, grasshopper)
Venn diagrams
webbings (mappings)



— list the directions in simple vocabulary on a chart: e.g.
1. Read the story (pages 4,5,6).
2. Go as fast or as slow as you need to.
3. Ask the teacher if you need help with the directions
4. Write your answers in the booklet.

— break down long sentences into shorter ones
— give more examples or different examples than the ones given (ones that are more relevant to your students)
— explain the way a multiple meaning is being used in this particular instance (especially if students would usually use it differently e.g. Exercise)
— simplify checklists and put on charts that the whole class can see easily
— give directions in the order that student will do the activities (saying "we will do....later, but first we will do this" is very confusing to LEP students)
— make connections with previous learning-
e.g. "Remember when we read_____. What kind of story did we say that was? Yes, it was a fable. The story you're going to read today is also a fable."
— when going to a new session, remind the students of vocabulary and information taught in the last session.

BEST COPY AVAILABLE

Reading, Writing and Math Assessments:

Reading Assessment:

1. For the directions, use strategies previously given.
2. Provide a regular dictionary, a picture dictionary and a thesaurus. (English or English/primary language)
3. Teachers **may not** give the meaning of a word, paraphrase or translate a word of the actual reading text of the reading assessments.

Writing and Math Assessments:

1. Teachers may use all of the strategies found in the section on directions and for the actual text of the assessments.)
2. Teachers may provide a regular dictionary, a picture dictionary and a thesaurus. (English or English/primary language)

*This list was approved by the ASAP Unit for all students, not just LEP students. The strategies will not hurt any student and will probably help more students than just your LEP students.

Other Concerns:

1. If students are taking an English ASAP assessment, but answer in Spanish, do not translate answers into English. Assessments answered in Spanish will count and will be graded by Spanish scorers.

2. Teachers may use their professional judgment in determining if migrant students are in their first or second year of ESL.

3. Limited English students on IEP's (due to language) follow the same guidelines as those for ESL students.

4. Paraprofessionals may assist limited English proficient students with ASAP under the direction of a certified teacher. Teachers and paraprofessionals may work as a team to determine which strategies would be appropriate for specific students according to the guidelines previously given.